Jeremiah Clemens-Dannecker

GEOG-435

Lab 03

Tutorial Questions

1. Refer back to the Spatial Autocorrelation report created in the "Quantify the significance of the aggregations" section. What was the distance threshold (selected by default) (*1 pt.*)?

The distance threshold was 58,968.9024 U.S. Feet

2. Select two other bandwidths, one reasonably greater than the default distance and one reasonably less than the default distance, and complete Spatial Autocorrelation using Moran's *I* again. What were the results based on each new bandwidth? Clustered, dispersed, or random (*1 pt.*)?

The first spatial autocorrelation I conducted was Clustered with a z-score of 7.3. The second spatial autocorrelation I conducted I chose a 30,000 feet bandwidth. Again the report states that the data is clustered and with a higher z score of 8.1. I then chose a bandwidth of 75,000 U.S. Feet and the data was still clustered with the lowest z score of 7.1

3. What other methods, besides nearest neighbor analysis, could be used to determine whether or not the spatial pattern is clustered, dispersed, or random *using the raw point data rather than the polygons* (*1 pt.*)?

You can use several different methods to determine the spatial pattern of raw point data.

Quadrat analysis - the study region is sampled using a set of quadrats. Counting the number of events in each can help determine whether the data is clustered, dispersed, or random.

Ripley's K-function - compares a given points distribution to a random distribution.

Kernel Density Estimation - estimates the probable density function of a variable representing the underlying distribution of the dataset.

4. Between DBSCAN (500 minimum features, search distance of 0.1 miles), HDBSCAN (500 minimum features) Multi-scale/OPTICS (500 minimum features, search distance of 0.1 miles), which two look the most similar (*1 pt.*)?

   The two that looked the most similar were the DBscan and the Multi-scale optics.

Supplementary Questions

5. Before completing any cluster analysis, use your best judgment - where does it appear that there are clusters of crashes involving cyclists (*1 pt.*)?

There is one big cluster in the center of Raleigh. I also see a cluster along Hillsborough St. west of Raleigh. I also see a few smaller clusters, especially northeast of Raleigh. There's one near the North Ridge Country Club, I see one on Capital Blvd. and a little west of Raleigh there's some along Western Blvd.

6. Test several methods of cluster analysis, then ultimately choose one that you believe represents the data well. State your method and parameters (*1 pt.*).

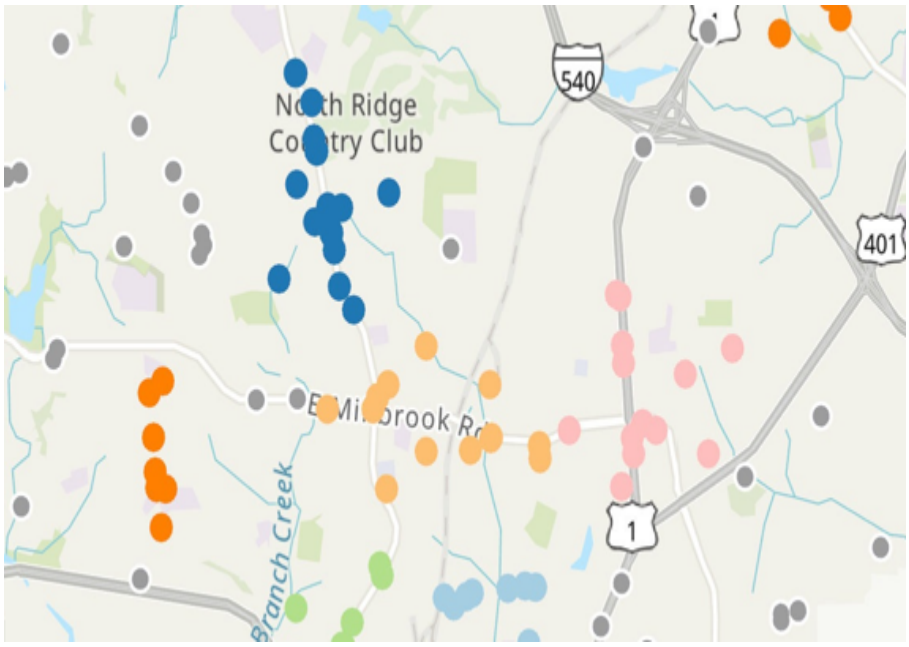
Of the three methods of cluster analysis I conducted I believe that the method that best represents the clusters in the data is DBSCAN with the minimum features set to 6 and the search distance set to 1000 meters.

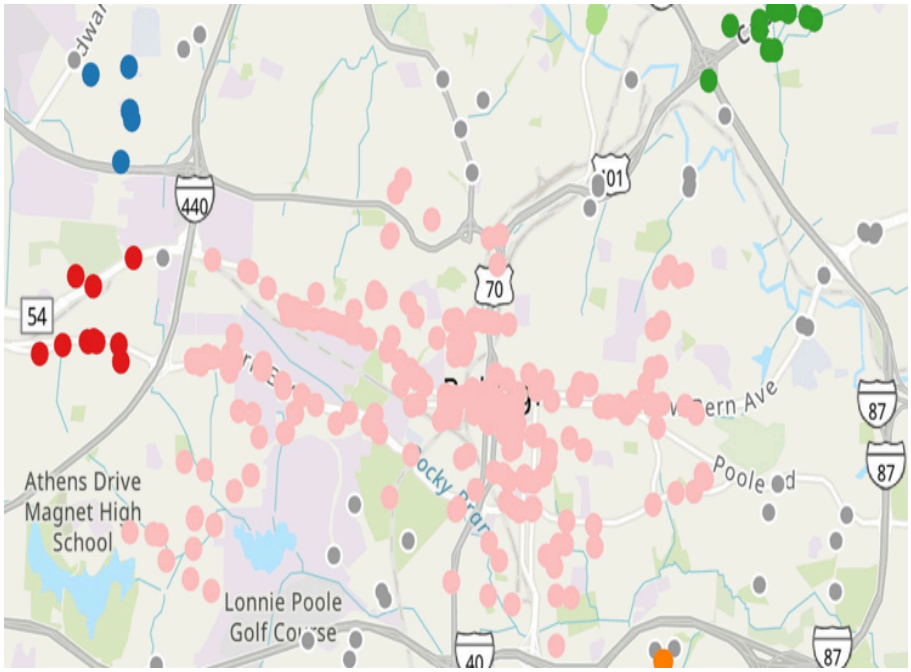
7. Create a map showing the significant clusters (*2 pts.*).

Map submitted into the Canvas Dropbox

8. How would you advise the city to act if they only can target two locations? Provide your reasoning (*2 pts.*).

   I would advise the city to take a look at these two areas in particular:





It is in these two areas of North Carolina, near the center of Raleigh that the city needs to target. These areas have by far the most amount of crashes compared to other areas. If these two areas get cleaned up, cyclists can feel much safer traveling.